

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **09026799 A**

(43) Date of publication of application: **28.01.97**

(51) Int. Cl.

G10L 3/00
G10L 3/00
G10L 3/00
G08G 1/0962
G09B 29/10
// G01C 21/00

(21) Application number: **07200362**

(22) Date of filing: **12.07.95**

(71) Applicant: **AQUEOUS RES:KK**

(72) Inventor: **YOKOYAMA SHOJI**
YAMAKAWA HIROYUKI
MURAKAMI YUMI

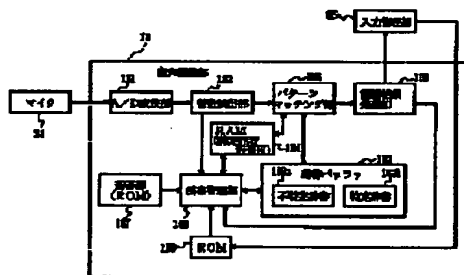
(54) **SPEECH RECOGNITION DEVICE**

COPYRIGHT: (C)1997,JPO

(57) Abstract:

PROBLEM TO BE SOLVED: To provide a speech recognition device which can efficiently use the storage capacity of a dictionary and is good in operability.

SOLUTION: When one operation is specified for the first time on a touch panel, etc., the standard pattern of the words indicating a specified operation is generated based on phoneme patterns representing features of respective phonemes stored in a dictionary part 167 and stored in a specific dictionary 163b in a dictionary buffer 163. A feature extraction part 162 extracts features from an input speech from a microphone 24 and generates a word pattern, and a pattern matching part 165 compares it with standard patterns stored in a RAM 164 and the dictionary buffer 163 to recognize the speech. Further, the word pattern generated by the feature extraction part 162 is stored as a standard pattern for a specific speaker in the RAM 164 and the specific dictionary 163b. Since a user registers standard patterns of only necessary words, an efficient dictionary is generated and an object of speech recognition can easily be understood.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-26799

(43) 公開日 平成9年(1997)1月28日

(51) Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/00	5 7 1		G 1 0 L 3/00	5 7 1 D 5 7 1 E 5 2 1 J C3 5 5 1 J C4
G 0 8 G 1/0962	5 2 1 5 5 1		G 0 8 G 1/0962	
審査請求 未請求 請求項の数4 F D (全 9 頁) 最終頁に続く				

(21) 出願番号 特願平7-200362

(22) 出願日 平成7年(1995)7月12日

(71) 出願人 591261509

株式会社エクス・リサーチ

東京都千代田区外神田2丁目19番12号

(72) 発明者 横山 昭二

東京都千代田区外神田2丁目19番12号 株

式会社エクス・リサーチ内

(72) 発明者 山川 博幸

東京都千代田区外神田2丁目19番12号 株

式会社エクス・リサーチ内

(72) 発明者 村上 ユミ

東京都千代田区外神田2丁目19番12号 株

式会社エクス・リサーチ内

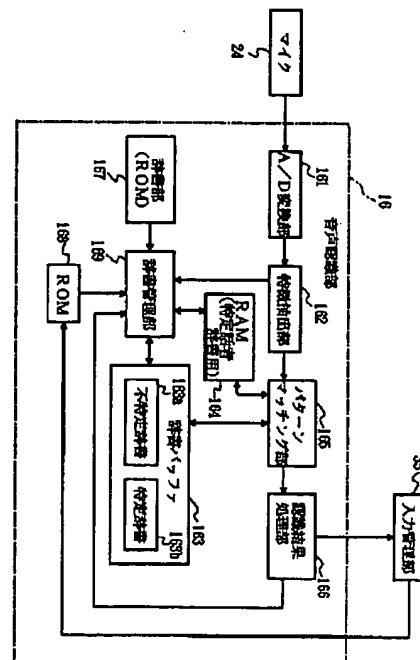
(74) 代理人 弁理士 川井 隆 (外1名)

(54) 【発明の名称】 音声認識装置

(57) 【要約】

【目的】 辞書の記憶容量を効率的に使用でき、操作性の良い音声認識装置を提供する。

【構成】 ある操作がタッチパネル等で初めて指定されると、指定操作を示す単語の標準パターンを、辞書部167に格納されている各音素の特徴を表す音素パターンに基づいて作成し、辞書バッファ163の特定辞書163bに格納する。そして、特徴抽出部162でマイク24からの入力音声の特徴を抽出して単語パターンが作成され、RAM164、辞書バッファ163に格納されている標準パターンとパターンマッチング部165で比較し、音声認識する。また、特徴抽出部162で作成した単語パターンを特定話者用標準パターンとしてRAM164および特定辞書163bに格納する。ユーザが必要な単語のみ標準パターン登録するので効率的な辞書が作成でき、音声認識対象も容易に理解できる。



【特許請求の範囲】

【請求項1】 所定の操作を選択する選択手段と、
この選択手段で選択可能な操作を表す単語を構成する各
音素についての特徴を表す音素パターンが予め格納され
た音素辞書と、

標準パターンが格納される辞書記憶手段と、

前記選択手段で選択された操作を表す単語の標準パター
ンを、前記音素辞書に格納された音素パターンから作成
する標準パターン作成手段と、

この標準パターン作成手段で作成された標準パターンを 10

前記辞書記憶手段に格納する不特定辞書格納手段と、

音声を入力する音声入力手段と、

この音声入力手段で入力された音声の特徴を抽出して対
応する単語パターンを作成する単語パターン作成手段
と、

この単語パターン作成手段で作成された単語パターンと
前記辞書記憶手段に格納された標準パターンとの類似度
から前記音声入力手段で入力された音声を認識する認識
手段と、

この認識手段による認識結果を出力する出力手段と、を 20
具備することを特徴とする音声認識装置。

【請求項2】 前記単語パターン作成手段で作成された
単語パターンを、特定話者用の標準パターンとして前記
辞書記憶手段に格納する特定辞書格納手段を具備するこ
とを特徴とする請求項1に記載の音声認識装置。

【請求項3】 前記辞書格納手段に格納された標準パ
ターンが所定数になった場合、前記不特定辞書格納手段お
よび前記特定辞書格納手段は、前記認識手段による認識
頻度が最も少ない標準パターンを削除することを特徴と
する請求項1または請求項2に記載の音声認識装置。 30

【請求項4】 ナビゲーション装置の入力手段として用
いたことを特徴とする請求項1、請求項2または請求項
3に記載の音声認識装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は音声認識装置に係り、詳
細には、特定の語句について発声された音声を認識する
音声認識装置に関する。

【0002】

【従来の技術】人間の話した音声を言葉として認識する
音声認識装置が各種方面で実用化されている。この音声
認識装置は、例えば、工場における各種装置に対応する
指示をはなれた場所から音声で指示する入力装置として
実用化されており、また、自動車のナビゲーション装置
において、目的地や指示情報等を音声入力する場合の音
声入力装置として用いることが考えられている。このよ
うな音声認識装置では、一般に入力された音声を特定す
るために、予め認識対象となる音声の周波数分布を分析
することで、例えば、スペクトルや基本周波数の時系列
情報等を特徴として抽出し、そのパターンを各単語に対 50

応させて格納する音声認識用辞書を備えている。この音
声認識辞書に格納される辞書としては、一般に不特定多
数の話者を対象とした音声の周波数分布に対して平均化
処理を行ったものをパターン（以下を標準パターンとす
る）化し、そのパターンを単語と共に登録する不特定話
者辞書及び、特定の話者を対象に入力音声のパターンを
単語と共に登録する特定話者辞書がある。

【0003】そして、認識すべき音声が入力される
と、入力された音声の周波数パターンと両辞書に格納さ
れた各単語のパターンをパターンマッチングにより比較
照合し、各単語に対する類似度を算出する。つぎに算出
された類似度が最も高い単語（パターンが最も近い単
語）を、入力された音声であると認識し、その単語を出
力するようにしている。つまり、入力された単語の周波
数分布のパターンがどの単語パターンに最もよく似てい
るかを調べることによって、入力音声を判定している。

【0004】

【発明が解決しようとする課題】しかし、現在の音声認
識装置では、不特定辞書及び特定辞書における標準パ
ターンの記憶容量に限界があり、語彙数が増すと、多くの
標準パターンを両辞書に蓄積できなかった。このよう
に、標準パターンの記憶容量に限界があるので、予め選
択された単語についての標準パターンのみが、不特定辞
書に登録されている。しかし、選択された登録単語はユ
ーザにとって必ずしも必要なものであるとは限らず、不
要な単語である場合には使用された記憶容量が無駄にな
っていた。また、辞書に予めどのような単語が登録され
ているか、ユーザにとっては、使用マニュアル等で登録
された単語を確認しない限りわからなかった。このため、
従来の音声認識装置では、ユーザの操作性の面で問題が
あった。

【0005】そこで、本発明は以上の課題を解決するた
めになされたもので、辞書の記憶容量を効率的に使用で
き、操作性の良い音声認識装置を提供することを目的と
する。

【0006】

【課題を解決するための手段】請求項1記載の発明で
は、所定の操作を選択する選択手段と、この選択手段で
選択可能な操作を表す単語を構成する各音素について
の特徴を表す音素パターンが予め格納された音素辞書と、
標準パターンが格納される辞書記憶手段と、前記選択手
段で選択された操作を表す単語の標準パターンを、前記
音素辞書に格納された音素パターンから作成する標準パ
ターン作成手段と、この標準パターン作成手段で作成さ
れた標準パターンを前記辞書記憶手段に格納する不特定
辞書格納手段と、音声を入力する音声入力手段と、この
音声入力手段で入力された音声の特徴を抽出して対応す
る単語パターンを作成する単語パターン作成手段と、こ
の単語パターン作成手段で作成された単語パターンと前
記辞書記憶手段に格納された標準パターンとの類似度か

ら前記音声入力手段で入力された音声を確認する認識手段と、この認識手段による認識結果を出力する出力手段と、を音声認識装置に具備させて前記目的を達成する。請求項2に記載の発明では、請求項1に記載の音声認識装置において、前記単語パターン作成手段で作成された単語パターンを、特定話者用の標準パターンとして前記辞書記憶手段に格納する特定辞書格納手段を、さらに具備させる。請求項3に記載の発明では、請求項1または請求項2に記載の音声認識装置において、前記辞書格納手段に格納された標準パターンが所定数になった場合、前記不特定辞書格納手段および前記特定辞書格納手段は、前記認識手段による認識頻度が最も少ない標準パターンを削除する。請求項4に記載の発明では、請求項1、請求項2、または請求項3に記載の音声認識装置を、ナビゲーション装置の入力手段として用いる。

【0007】

【作用】請求項1に記載の音声認識装置では、では、選択手段で選択可能な操作を表す単語を構成する各音素についての特徴を表す音素パターンを予め音素辞書に格納する。そして、選択手段で選択された操作を表す単語の標準パターンを音素パターンから作成し、辞書記憶手段に格納する。一方、音声入力手段で入力された音声の特徴を抽出して対応する単語パターンを作成し、辞書記憶手段に格納された標準パターンとの類似度から入力された音声を確認して、認識結果を出力する。請求項2に記載の音声認識装置では、入力された音声から作成された単語パターンを特定話者用の標準パターンとして辞書記憶手段に格納する。請求項3に記載の音声認識装置では、辞書格納手段に格納された標準パターンが所定数になった場合、認識頻度が最も少ない標準パターンを削除する。請求項4に記載の音声認識装置では、ナビゲーション装置の入力手段として用いる。

【0008】

【実施例】以下、本発明の音声認識装置における一実施例を図1ないし図4を参照して詳細に説明する。

①実施例の概要

本実施例の音声認識装置では、ある操作がタッチパネル等で初めて指定された時点で、指定された操作を示す単語の標準パターンを、各音素の特徴を表す音素パターンに基づいて作成し、不特定話者用の辞書に格納する。音声が入力されると、入力音声の特徴を抽出して単語パターンを作成し、辞書に格納されている標準パターンとの類似度から音声認識する。この音声認識の際に作成した単語パターンを特定話者用の標準パターンとして辞書に格納する。このように本実施例では、一度タッチパネル等から指定された必要な単語についてのみ、標準パターンが作成され、次回から音声で入力することができるようになる。音声認識の対象となる単語は、一度ユーザによって指定入力された単語なので、ユーザは認識対象を改めて確認する必要がなくなる。

【0009】②実施例の詳細

図1は本発明の一実施例に係る音声認識装置をナビゲーション装置に適用した場合のシステム構成を表したものである。このナビゲーション装置は、演算部10を備えている。この演算部10には、タッチパネルとして機能するディスプレイ11aとこのディスプレイ11aの周囲に設けられた操作用のスイッチ11bとを含む表示部11と、この表示部11のタッチパネルやスイッチ11bからの入力を管理するスイッチ入力類管理部12が接続されている。

【0010】スイッチ11bには、ナビゲーションのメニュー画面を指定するスイッチ、エアコンの調整用のスイッチ、オーディオの操作を行うためのスイッチ等の各種スイッチがある。これらのスイッチを押すと、対応するメニュー画面がディスプレイ11aに表示されるようになっていく。タッチパネル11aに表示される画面は、階層構造になっており、最上位層にメニュー画面がある。そして、ナビゲーション用のメニュー画面には、例えば、目的地設定や、地名検索等を指定する指定キーが表示され、目的地設定キーが指定されると、更にその下位層の画面としてスキー場、ゴルフ場等を指定する指定キーが表示されるようになっており、各指定キーの指定によって最下層の画面まで順次表示されるようになっていく。本実施例では、これらタッチパネル11aの画面に表示される各種指定キー、およびスイッチ11bの各種スイッチの内容が、音声認識の対象となる。すなわち、これら各種の指定キーとスイッチが最初に押下される毎に、音声認識のための不特定話者辞書が作成され、それ以後、作成された辞書に基づいて、入力された音声の認識が行われると共に、入力音声による特定話者辞書が作成されるようになっていく。

【0011】また、演算部10には、現在位置測定部13と、速度センサ14と、地図情報記憶部15と、本実施例における音声認識部16と、音声出力部17とが接続されている。現在位置測定部13は、緯度と経度による座標データを検出することで、車両が現在走行または停止している現在位置を検出する。この現在位置測定部13には、人工衛星を利用して車両の位置を測定するGPS(Global Position System)レシーバ21と、路上に配置されたビーコンからの位置情報を受信するビーコン受信装置20と、方位センサ22と、距離センサ23とが接続され、現在位置測定部13はこれらからの情報を用いて車両の現在位置を測定するようになっていく。

【0012】方位センサ22は、例えば、地磁気を検出して車両の方位を求める地磁気センサ、車両の回転角速度を検出しその角速度を積分して車両の方位を求めるガスレートジャイロや光ファイバジャイロ等のジャイロ、左右の車輪センサを配置しその出力パルス差(移動距離の差)により車両の旋回を検出することで方位の変位量を算出するようにした車輪センサ、等が使用される。距

離センサ23は、例えば、車輪の回転数を検出して計数し、または加速度を検出して2回積分するもの等の各種の方法が使用される。なお、GPSレシーバ21とビーコン受信装置20は単独で位置測定が可能であるが、GPSレシーバ21やビーコン受信装置20による受信が不可能な場所では、方位センサ22と距離センサ23の双方を用いた推測航法によって現在位置を検出するようになっている。

【0013】地図情報記憶部15は、例えばCDROM等の大容量記憶装置で構成されている。この地図情報記憶部15には、目的地までの経路探索に必要な道路データや、探索した経路をディスプレイ11aに表示するための地図データ等の、経路探索および経路案内に必要な各種データが格納されている。音声認識部16には、人間の音声や、電話番号に対応した発信音が入力されるマイク24が接続されている。音声出力部17は、音声を電気信号として出力する音声出力用IC26と、この音声出力用IC26の出力をディジタル-アナログ変換するD/Aコンバータ27と、変換されたアナログ信号を増幅するアンプ28とを備えている。アンプ28の出力端にはスピーカ29が接続されている。

【0014】演算部10は、CPU（中央処理装置）、ROM（リード・オンリ・メモリ）、RAM（ランダム・アクセス・メモリ）等を備え、CPUがRAMをワーキングエリアとしてROMに格納されたプログラムを実行することによって、上記の各構成を実現するようになっている。すなわち、演算部10は、速度センサ14および地図情報記憶部15に接続された地図データ読込部31と、地図描画部32と、地図データ読込部31および地図描画部32を管理する地図管理部33と、地図描画部32および表示部11に接続された画面管理部34と、スイッチ入力類管理部12および音声認識部16に接続された入力管理部35と、音声出力部17の音声出力用IC26に接続された音声出力管理部36、および、地図管理部33、画面管理部34、入力管理部35、音声出力管理部36を管理する全体管理部37とを備えている。入力管理部35は、ディスプレイ11aのタッチパネルに表示される各種指定キーや、スイッチ11bが初めて選択されたものか否かも管理するようになっている。そして、スイッチ等の選択が初めて行われたものである場合、音声認識部16に、そのスイッチ等に応じたタッチパネル等入力信号を入力管理部35に供給するようになっている。

【0015】図2は、図1における音声認識部16の構成を示すブロック図である。この図に示すように、音声認識部16は、マイク24から入力される音声信号をディジタル信号に変換するA/D変換部161と、このD/A変換部161の出力信号から入力された音声についての特徴を抽出して対応する単語パターンを作成する特徴抽出部162と、所定の音声にする標準パターンが格

納される辞書としての辞書バッファ163とRAM164とを備えている。辞書バッファ163は、不特定話者認識用として一般的な標準パターンが格納される不特定辞書163aと、マイク24から入力された特定話者の音声から作成された単語パターンが標準パターンとして格納される特定辞書163bを備えている。ここで、標準パターンとは、所定単位毎での音声信号についての、スペクトルや基本周波数の時系列情報である。所定単位毎の音声信号としては、音節、単語、音素、半音節、単語間、音素間、音節間、半音節間等の単位毎の音声信号が使用される。また、特徴抽出部162は、多チャネル・バンドパスフィルタや線形予測分析等によって、この単語パターンを抽出するようになっている。

【0016】また、音声認識部16は、辞書バッファ163とRAM164に格納された標準パターンと、特徴抽出部162によって抽出された単語パターンとを比較するパターンマッチング部165と、パターンマッチング部165の比較結果に基づいてマイク24から入力された音声の内容を認識し、その認識内容に従って、演算部10に対する音声入力信号を生成して、演算部10の入力管理部35へ出力する認識結果処理部166とを備えている。

【0017】さらに、音声認識部16は、辞書部167と、ROM168、および辞書管理部169を備えている。辞書部167はROMで構成され、各音素の特徴を表す音素パターンが格納されている。この音素パターンは、不特定辞書を辞書管理部169で作成するためのもので、複数人のアナウンサー等による発声音を平均化することで各音素毎に作成されたものが格納されている。なお、辞書部167のROMは各種記憶装置を使用することが可能であるが、CD（コンパクトディスク）ROMを使用するようによい。

【0018】ROM168には、ユーザによって押下された、ディスプレイ11aのタッチパネルに表示された指定キーや、スイッチ11bのスイッチに応じたタッチパネル等入力信号が入力管理部35から供給されるようになっている。そして、ROM168は、タッチパネル等入力信号に応じた単語を示す音素指定信号に変換して辞書管理部169に供給するようになり、そのための変換テーブルを備えている。例えば、ディスプレイ11aのタッチパネルから「目的地設定」の指定キーが押下され、それに対応するタッチパネル等入力信号が入力管理部35から供給された場合、ROM168は、タッチパネル等入力信号に対応した単語「もくてきちせつてい」を示す音素指定信号に変換して辞書管理部169に供給する。

【0019】辞書管理部169は、ROM168から供給された音素指定信号に対応する各音素パターンを辞書部167から読み出して、不特定話者認識用の辞書を合成し、これを辞書バッファ163の不特定辞書163a

に格納するようになっている。また辞書管理部169には、マイク24から入力された音声についての認識が成功すると、認識結果処理部166から音声入力信号が供給されるようになっている。辞書管理部169は、この音声入力信号が供給されると、その認識された音声について特徴抽出部162で抽出した単語パターンを標準パターンとして、特定辞書163bとRAM164に格納するようになっている。ここで、RAM164は、表示部11で押下可能なスイッチと指定キーの数だけ(=音声入力信号の数だけ)の単語を格納するエリアが確保されており、各エリアには1つの標準パターンが格納される。従って、マイク24から入力された音声認識される毎に、その入力音声に対応する格納エリアが、新しい標準パターンによって更新されるようになっている。また、RAM164には、入力音声の音素パターンも音素辞書として更新される。すなわち、マイク24から入力された音声認識される毎に、その入力された音声について特徴抽出部162で抽出した単語パターンから各音素毎の音素パターンを作成し、RAM164が更新される。

【0020】また特定辞書163bにも、音声入力信号の数だけの単語を格納するエリアが確保されており、各エリアには、複数の標準パターンが格納されるようになっている。従って、音声認識される毎に、その入力された音声に対応する格納エリアに新しい標準パターンが順次蓄積されるようになっている。なお、音声入力信号に対応する格納エリアに、所定数の標準パターンが格納された以後は、最も古い標準パターンと最新の標準パターンとの間で更新が行われる。

【0021】次に、このように構成された実施例の動作について説明する。図3は、ユーザによるナビゲーション装置の使用状態を表したものである。本実施例のナビゲーション装置は、システムの初期状態において、RAM164と辞書バッファ163の不定辞書163a、特定辞書163bに、標準パターンが格納されていない。従って、システムの初期状態では音声認識を行うことができず、図3(a)~(d)に示すように、順次表示装置11のスイッチ11bおよびディスプレイ11aのタッチパネルから所定の処理を選択する。ここで、ナビゲーション用のメニュー画面(a)から、タッチパネルで順次目的地の入力を行う場合について説明する。まず、メニュー画面(a)において、ユーザが「目的地設定」を選択すると、その下位層の目的地設定画面(b)が表示される。同様に、目的地設定画面(b)において「ゴルフ場」を選択し、さらに、「千葉県」、「習志野CC」というように順次選択することで、全体管理部37により、習志野カントリークラブが目的地として設定される。

【0022】そして、表示部11のディスプレイ11aのタッチパネルやスイッチ11bのスイッチが、初めて

選択される毎に、そのスイッチに対応する単語の標準パターンが辞書管理部169で作成され、辞書バッファ163の不定辞書163に格納される。すなわち、図3(a)~(d)の順に選択が行われると、選択に対応する単語「目的地設定」「ゴルフ場」、「千葉県」、「習志野カントリークラブ」の標準パターンが順次特定辞書163aに格納される。

【0023】このようにして、表示部11の操作によって目的地設定が行われ、同時に、対応する指定キーやスイッチが意味する単語の不定辞書が作成されると、それ以後は、音声による入力が可能になる。すなわち、図3(a)に示すナビゲーション用のメニュー画面において、(e)に示すようにユーザが「目的地設定」と発声してマイク24に入力すると、音声認識部16で認識され、ディスプレイ11aには目的地設定画面(b)を表示する。更に、ユーザが「ゴルフ場」と発声すると(図3(f))、ゴルフ場画面が表示される。同様に、ユーザが「千葉県」、「習志野カントリークラブ」と順に発声すると、その音声認識され、最終的に地習志野CCが目的地として設定される。

【0024】ここで、ユーザが発声し認識された音声については、その音声の単語パターンが特定話者用の標準パターンとして、特定辞書136bとRAM164に格納される。なお、図3(g)において、ユーザが「千葉県」と発声せずに、「北海道」と発声した場合、北海道という単語の不定辞書はまだ作成されていないので、認識することができない。音声「北海道」についての認識を可能にするためには、図3(c)のゴルフ場画面または他の画面に表示されている「北海道」のキーを一度選択する必要がある。

【0025】次に動作の詳細について説明する。図4は、辞書作成および作成した辞書による音声認識の動作を表したフローチャートである。まず、制御部10は、入力管理部35にデータが入力されたか否かを判断し、入力があった場合(ステップ31;Y)、入力されたデータがディスプレイ11のタッチパネルまたはスイッチ11bからの入力か否かを判断する(ステップ32)。

【0026】タッチパネル等からの入力である場合(ステップ32;Y)、入力管理部35は、入力されたキーの使用が初めてか否かを判断することで、選択対象となったキーを表す単語の標準パターンが不定辞書にあるか否かを判断する(ステップ33)。選択対象の不定辞書がない場合(ステップ33;Y)、音声認識部16は、選択対象の不定辞書を作成する(ステップ34)。すなわち、音声認識部16のROM168は、入力管理部35から供給されるタッチパネル等入力信号を、対応する単語を示す音素指定信号に変換して辞書管理部169に供給する。辞書管理部169は、供給された音素指定信号により辞書部167から各音素の音素パターンを読み出し、ステップ32で入力されたキーを表

す単語に対する不特定話者用辞書を作成して不特定辞書163に格納する。

【0027】ステップ34において不特定辞書が作成された後、または、選択対象の不特定辞書がある場合(ステップ33;N)、入力管理部35は、タッチパネル等からの入力を全体管理部37に供給する(ステップ35)。全体管理部37は、タッチパネル等からの入力に応じて、画面管理部34を制御し、つぎ認識対象語彙を選択するための画面に切り換えて(ステップ36)、メインルーチンにリターンする。

【0028】一方、ステップ32において、タッチパネル等からの入力ではなく音声入力である場合(;Y)、音声認識部16は、ユーザによって発声された音声をマイク24から入力し(ステップ37)、入力音声についての音声認識を行う(ステップ38)。すなわち、音声認識部16は、入力された音声データをA/D変換部161でデジタルデータに変換し、特徴抽出部162で入力音声の単語パターンを抽出して、パターンマッチング部165と辞書管理部169に供給する。

【0029】パターンマッチング部165では、入力音声の単語パターンと、まず、RAM164に格納されている各単語についての最新の標準パターンとを比較し、各単語との類似度を算出し、類似度が最も高い単語を入力された音声であると認識する。いずれの単語との類似度も、所定の閾値以下であった場合には認識不能であるため、次に特定辞書163bに格納されている各単語の標準パターンと比較して類似度を算出し、閾値よりも大きくて最も類似度が高い単語を音声であると認識する。特定辞書163bの標準パターンとの類似度もすべて閾値以下である場合には、さらに、特定辞書163aに格納されている、不特定話者用の全標準パターンと比較し、閾値よりも大きくて最も類似度が高い単語を音声であると認識し、いずれも閾値以下である場合には、入力された音声について認識不可能であると判断する。パターンマッチング部165による認識結果がでると、認識結果処理部166では、認識内容に従って音声入力信号を生成し、演算部10の入力管理部35に供給する。

【0030】そして、制御部10の全体管理部37は、入力管理部35に供給された認識結果が認識不能である場合(ステップ39;N)、画面管理部34を制御して、入力された音声に対応する辞書が存在しない旨の表示をディスプレイ11aに表示して(ステップ40)、メインルーチンにリターンする。一方、標準パターンとの類似度が閾値よりも大きく、入力音声の認識が成功した場合(ステップ39;Y)、特定辞書の更新を行う(ステップ41)。すなわち、認識結果処理部166は認識した単語に対応する音声入力信号を辞書管理部169にも供給する。辞書管理部169では、音声入力信号が供給されると、RAM164のその音声入力信号に対応した格納エリアを、特徴抽出部162から供給された

単語パターンに更新する。更に、辞書管理部169は、特徴抽出部162から供給された単語パターンを特定辞書163bの対応する格納エリアに格納するか、または、その格納エリアに格納されている最も古い標準パターンを供給された単語パターンに書き換える。

【0031】ステップ41において、特定辞書の更新が終了すると、ステップ35に移行し、全体管理部37は、マイク24から入力され音声認識部16で認識された音声の音声入力信号を、全体管理部37に供給する。そして、全体管理部37は、音声入力信号に応じて、画面管理部34を制御し、つぎ認識対象語彙を選択するための画面に切り換えて(ステップ36)、メインルーチンにリターンする。

【0032】なお、以上説明した実施例では、RAM164と不特定辞書163aおよび特定辞書163bは、表示部11で押下可能なスイッチと指定キーの数だけの単語を格納するエリアが確保されていたが、これよりも少ない数、例えば半分や1/3の格納エリアとしてもよい。この場合、辞書管理部169は、エリア不足になった場合に対応して、各エリアに格納された標準パターンが示す単語の使用頻度を計数するようにしておき、最も使用頻度が少ない格納エリアをクリアし、新しい単語パターン、または標準パターンの入力用に確保するようにする。これにより、認識率の低下を抑えつつ、記憶容量を減らすことができる。また、パターンマッチング分165で比較する辞書数が減少するので、認識速度も向上させることができる。また、認識語彙数や辞書バッファ等の記憶容量に制約がある場合でも、使用によって辞書の再構築を行うことで効率の良いシステムとすることができ。

【0033】また、ROM168から出力される音素指定信号を出力したが、その音素指定信号としてテキスト文字列を示すデータを使用してもよい。

【0034】また、その音声の入力によって表示される画面を特定する画面ポインタを、認識音声に対する標準パターンの各々に対応付けて記憶することで、ダイレクトに画面を呼び出すことができるようにしてもよい。例えば、「目的地設定」という音声で認識された場合に対応する画面として図3(b)が表示される場合には、音声「目的地設定」の標準パターンに対応付けて図3

(b)を特定する画面ポインタが併せて記憶される。同様に、音声「ゴルフ場」の標準パターンに対応付けて図3(c)の画面ポインタが、音声「千葉県」の標準パターンに対応付けて図3(d)の画面ポインタが併せて記憶される。以後は、目的地設定、ゴルフ場、千葉県、習志野CCの順に音声の入力と認識を行わなくても、「ゴルフ場」と音声入力することで、画面ポインタから図3(c)の画面が表示される。その後、「習志野CC」と音声入力することで、習志野CCが目的地として設定され、入力処理が容易になる。さらに、音声の標準パター

11

12

ンに画面を特定する画面ポインタを併せて記憶するのではなく、音声の標準パターンに、一連の入力処理手順や入力履歴を併せて格納するようにしてもよい。これによって、図3の例の場合、直接「習志野CC」と音声を入力するだけで目的地設定が終了するようになり、より入力処理を容易かつ迅速に行うことができる。

【0035】

【発明の効果】本発明の音声認識装置によれば、辞書の記憶容量を効率的に使用でき、操作性を向上させることができる。

【図面の簡単な説明】

【図1】本発明の一実施例に係る音声認識装置をナビゲーション装置に適用した場合のシステム構成図である。

【図2】同上、音声認識部の構成を示すブロック図である。

【図3】同上、ユーザによるナビゲーション装置の使用状態を表す説明図である。

【図4】同上、辞書作成および作成した辞書による音声認識の動作を表したフローチャートである。

【符号の説明】

10 演算部

* 11 表示部

11a ディスプレイ

13 現在位置測定部

15 地図情報記憶部

16 音声認識部

161 A/D変換部

162 特徴抽出部

163 辞書バッファ

164 RAM

10 165 パターンマッチング部

166 認識結果処理部

167 辞書部

168 ROM

169 辞書管理部

17 音声出力部

24 マイク

33 地図管理部

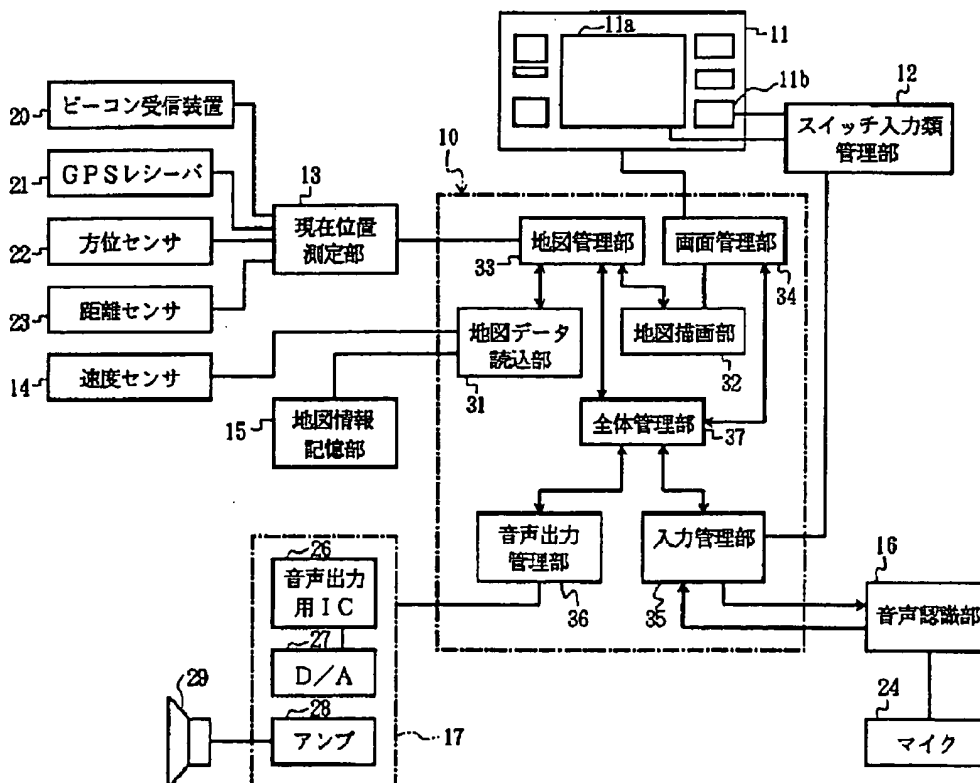
34 画面管理部

35 入力管理部

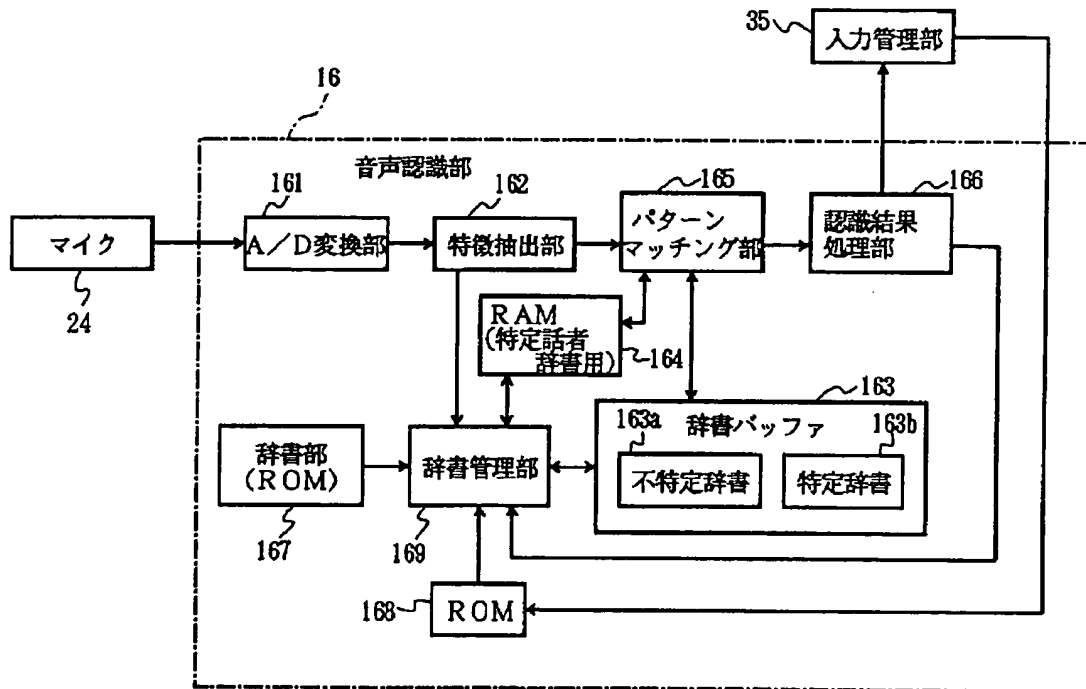
20 37 全体管理部

*

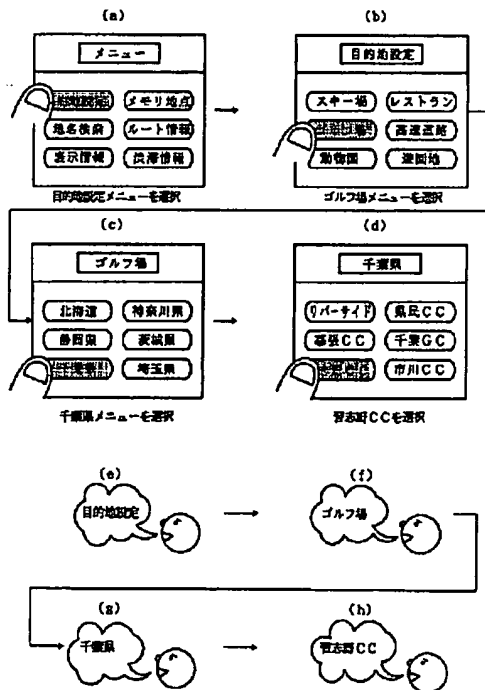
【図1】



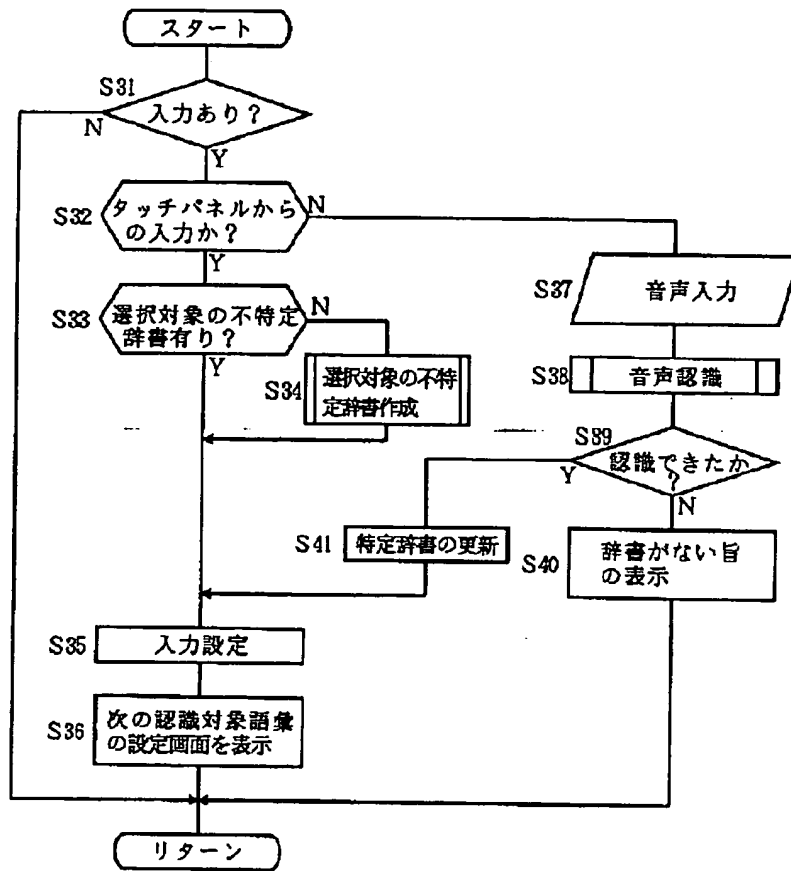
【図2】



【図3】



〔図4〕



フロントページの続き

(51)Int.Cl.⁶

G 0 9 B 29/10

// G 0 1 C 21/00

識別記号

庁内整理番号

F I

G 0 9 B 29/10

G 0 1 C 21/00

技術表示箇所

A

H